MAXIMUM LIKELIHOOD ESTIMATION OF STABILITY PARAMETERS FOR THE STANDARD GYROSCOPE ERROR MODEL

Reid Reynolds
Spectrum Astro, Inc.

ABSTRACT

Stability parameters for the standard gyroscope error model typically are estimated using subjective or non-optimal methods which do not provide an objective measure of statistical uncertainty. In this paper, we present a set of equations for maximum likelihood estimation of stability parameters which, as typical of maximum likelihood estimation procedures, automatically provides an uncertainty measure in the form of the estimated Fisher information matrix. Since the computational burden associated with the equations is daunting even by the standard of today's computational resources, a sub-optimal algorithm is proposed which asymptotically approaches the accuracy of the optimal algorithm and, produces estimates of parameter uncertainties as well.

INTRODUCTION

Accurate estimates of gyroscope stability parameters are required to obtain optimal performance from attitude filters that combine measurements from absolute attitude sensors with the attitude propagated from gyroscope data (ref. 1). Some analytical techniques which have been employed to estimate these parameters include Power Spectral Density (PSD) estimation (ref. 2 & 3) and Allan Variance (ref. 4). While these graphical techniques, particularly the PSD, are well suited for qualitative analysis of gyro data, they remain significantly subjective and, there is no readily accessible quantifier of the accuracy of the estimates they provide.

In 1980, Wyman and Sargent of the TRW Space and Electronics Group created the AUTOFIT program (ref. 5) in an effort to remove some of the subjectivity from the estimation process and create a uniform industry standard by which the competing gyroscopes could be judged. AUTOFIT is robust and easy to use and a good standard for assessing gyro performance. However, even AUTOFIT is still not optimal and does not provide an estimate of uncertainty in the stability parameters it provides.

This paper documents a new maximum likelihood approach (ref. 6) for estimating gyro stability parameters. While the approach is conceptually simple, it requires a vast amount of computer operations to implement as a truly optimal algorithm. Hence, we present a suboptimal algorithm which asymptotically approaches the accuracy of the optimal algorithm. Like all maximum likelihood estimators, these algorithms have the significant advantage that they produce estimates of parameter uncertainties in the form of the estimated Fisher information matrix.

The maximum likelihood algorithm is still vulnerable to data which does not conform to the Standard Gyroscope Error Model (SGEM) or, is tainted by periodic or other disturbances. For identifying such disturbances, there is probably no better tool than a Fast Fourier Transform (FFT) based PSD. Plotting the computed Allan Variance against that predicted by the estimated stability parameters also provides a good sanity check on the final product, as well as giving a direct visual indicator of optimal gyro averaging time. Thus, a good program for estimating gyro stability parameters is to use all of the available tools, relying upon each in its area of strength.

STANDARD GYROSCOPE ERROR MODEL

The Standard Gyroscope Error Model or SGEM is depicted in Figure 1 below. The final output is average gyro rate error over the interval *T*. The process m_k represents measurement noise, usually (but not necessarily) dominated by quantization error, modeled as a sequence of independent samples with uniform probability density in the interval $[-Q/2, Q/2]$, where Q is the quantization interval and $\sigma_m = Q/\sqrt{12}$. The measurement noise produces an error in angular readout hence, this process is often referred to as *Angle White Noise (AWN).*

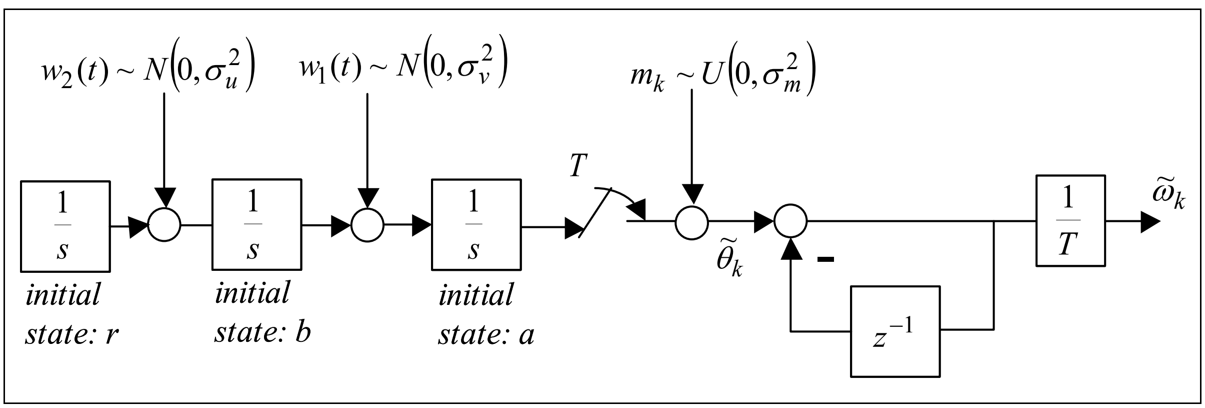

Figure 1. Standard Gyroscope Error Model

The first integrator prior to the sampler is driven in part by a wideband noise process $w_1(t)$, modeled as zero mean white noise with Normal density and spectral density of σ_v^2 in units of angle squared per unit time. The result is a Wiener process which, upon sampling, becomes a random walk in angle. Hence, the parameter σ_v is often informally referred to as the *Angle Random Walk (ARW)* parameter (the standard deviation of the resulting random walk in angle is this parameter multiplied by the squareroot of time).

In similar fashion, the noise process $w_2(t)$, quantified by the parameter σ_u (in units of angle per time$^{3/2}$), produces a random walk in rate and, this parameter is often referred to as the *Rate Random Walk (RRW)* parameter. The output state of the second integrator is actually the dynamically changing gyro bias at the given instant of time.

The third integrator is usually not modeled in attitude filter mechanizations because its effects are usually not observable within any reasonable time span relative to the filter time constants. However, over long test intervals, a ramp error in rate can become clearly discernible.

The autocorrelation of the angular error $\tilde{\theta}_k$ is

$$R_\Theta(k_1,k_2) = \left(a + bTk_1 + \frac{1}{2} r T^2 k_1^2\right)\left(a + bTk_2 + \frac{1}{2} r T^2 k_2^2\right) +$$
$$\sigma_m^2 \delta_{k_1 k_2} + \sigma_v^2 T \min(k_1,k_2) + \frac{\sigma_u^2 T^3}{6} \min(k_1^2,k_2^2)\left(3\max(k_1,k_2) - \min(k_1,k_2)\right) \tag{1}$$

The autocorrelation of $\tilde{\omega}_k$ is therefore

$$R_\Omega(k_1,k_2) = \left(b + rT(k_1 - 1/2)\right)\left(b + rT(k_2 - 1/2)\right) +$$
$$\frac{\sigma_m^2}{T^2}\left(2\delta_{k_1 k_2} - \delta_{k_1,k_2-1} - \delta_{k_1-1,k_2}\right) + \left(\frac{\sigma_v^2}{T} - \frac{\sigma_u^2 T}{6}\right)\delta_{k_1 k_2} + \sigma_u^2 T \min(k_1 - 1/2, k_2 - 1/2) \tag{2}$$

Another error process which occasionally appears in gyro data is the mysterious *flicker noise*, also called *1/f noise* because of the -1 decade per decade slope it produces on a log-log plot of the PSD. Such frequency dependent noise can (but does not necessarily) result as average behavior of processes governed by partial differential equations (PDE's). Such PDE's generally can be solved via modal decomposition into an infinite sum of complex exponential processes and, such models can sometimes be truncated to provide an acceptable finite dimensional model for a given *1/f* noise process. However, such modeling is difficult and time consuming and does not lend itself to be easily codified into a uniform standard for gyroscope error modeling. As such, we will not address *flicker noise* here and, our best advice is, if you see it in your data, advise the gyro manufacturers to try to isolate the cause and get rid of it.

MAXIMUM LIKELIHOOD ESTIMATION OF SGEM PARAMETERS

A method for achieving maximum likelihood estimation of the stability parameters is as follows. First, we would like to separate the estimation of the statistical parameters σ_m, σ_v, and σ_u, from the estimation of the random constant ramp-in-rate-parameter r. One way to do this is to difference the angle data thrice i.e., let

$$\alpha_k = \frac{\tilde{\omega}_k - \tilde{\omega}_{k-1}}{T} \tag{3}$$

$$\beta_k = \frac{\alpha_k - \alpha_{k-1}}{T} \tag{4}$$

From equation (2), it is not difficult to find that β_k is a zero mean process with autocorrelation

$$
\begin{aligned}
R_B(k_1,k_2) = \frac{\sigma_m^2}{T^6}&\left(20\delta_{k_1k_2} - 15\delta_{k_1,k_2-1} - 15\delta_{k_1-1,k_2} + 6\delta_{k_1,k_2-2} + 6\delta_{k_1-2,k_2} - \delta_{k_1,k_2-3} - \delta_{k_1-3,k_2}\right) + \\
&\frac{\sigma_v^2}{T^5}\left(6\delta_{k_1k_2} - 4\delta_{k_1,k_2-1} - 4\delta_{k_1-1,k_2} + \delta_{k_1,k_2-2} + \delta_{k_1-2,k_2}\right) + \\
&\frac{\sigma_u^2}{6T^3}\left(6\delta_{k_1k_2} - 2\delta_{k_1,k_2-1} - 2\delta_{k_1-1,k_2} - \delta_{k_1,k_2-2} - \delta_{k_1-2,k_2}\right)
\end{aligned} \tag{5}
$$

For a set $\{\beta_k \mid k = 1, N\}$ of random variables, construct the $N \times N$ Toeplitz matrices

$$
\mathbf{P}_1 = \begin{bmatrix}
20 & -15 & 6 & -1 & & & & & & \\
-15 & 20 & -15 & 6 & -1 & & & & & \\
6 & -15 & 20 & -15 & 6 & -1 & & & & \\
-1 & 6 & -15 & 20 & -15 & 6 & -1 & & & \\
 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \\
 & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\
 & & & -1 & 6 & -15 & 20 & -15 & 6 & -1 \\
 & & & & -1 & 6 & -15 & 20 & -15 & 6 \\
 & & & & & -1 & 6 & -15 & 20 & -15 \\
 & & & & & & -1 & 6 & -15 & 20
\end{bmatrix}
$$

$$
\mathbf{P}_2 = \begin{bmatrix}
6 & -4 & 1 & & & \\
-4 & 6 & -4 & 1 & & \\
1 & -4 & 6 & -4 & 1 & \\
 & \ddots & \ddots & \ddots & \ddots & \ddots \\
 & & 1 & -4 & 6 & -4 \\
 & & & 1 & -4 & 6
\end{bmatrix}
$$

$$
\mathbf{P}_3 = \frac{1}{6}\begin{bmatrix}
6 & -2 & -1 & & & \\
-2 & 6 & -2 & -1 & & \\
-1 & -2 & 6 & -2 & -1 & \\
 & \ddots & \ddots & \ddots & \ddots & \ddots \\
 & & -1 & -2 & 6 & -2 \\
 & & & -1 & -2 & 6
\end{bmatrix}
$$

and let $\mathbf{P} = \frac{\sigma_m^2}{T^6}\mathbf{P}_1 + \frac{\sigma_v^2}{T^5}\mathbf{P}_2 + \frac{\sigma_u^2}{T^3}\mathbf{P}_3$. Let $\underline{\beta}$ be the column *N* vector with elements equal to the values β_k. For the SGEM, we have allowed that the distribution of the measurement error may be mostly uniform due to quantization error. However, since each β_k is the sum of four other random variables, under the Central Limit Theorem, it is reasonable to approximate the distribution as jointly normal with probability distribution function (PDF)

$$f_B(\underline{\beta}) = \frac{1}{(2\pi)^{N/2}\sqrt{|\mathbf{P}|}} e^{-\frac{1}{2}\underline{\beta}^T \mathbf{P}^{-1} \underline{\beta}} \tag{6}$$

The method of *Maximum Likelihood* seeks to find the parameters σ_m, σ_v, and σ_u which maximize equation (6) for a given set of data $\{\beta_k \mid k = 1, N\}$. Since the *log* function is monotonic, maximizing (6) is the same as minimizing the negative *log-likelihood* function

$$-\log(f_B(\underline{\beta})) = \ \ldots \ + \frac{1}{2}\log|\mathbf{P}| \ + \ \frac{1}{2}\underline{\beta}^T \mathbf{P}^{-1} \underline{\beta} \tag{7}$$

You also get, as a byproduct, an estimate of the variance of your estimates in the form of the *Fisher Information Matrix* (ref. 6), which is the inverse of the expectation of the Hessian matrix (matrix of 2nd derivatives) of the negative log-likelihood function in equation (7).

Define the vector $\underline{\lambda} = \left[\sigma_m^2 / T^6 \quad \sigma_v^2 / T^5 \quad \sigma_u^2 / T^3\right]$. The first order condition for a minimum is that the gradient of the negative log-likelihood function must be zero. We need the following lemmas:

Lemma 1 – derivative of the inverse

The derivative of the inverse of a matrix $\mathbf{P}$ with respect to a scalar parameter γ is

$$\frac{\partial \mathbf{P}^{-1}}{\partial \gamma} = -\mathbf{P}^{-1}\frac{\partial \mathbf{P}}{\partial \gamma}\mathbf{P}^{-1} \tag{8}$$

Lemma 2 – derivative of the logarithm of the determinant

The derivative of the logarithm of the determinant of a matrix $\mathbf{P}$ with respect to a scalar parameter γ is

$$\frac{\partial \log|\mathbf{P}|}{\partial \gamma} = trace\left(\mathbf{P}^{-1}\frac{\partial \mathbf{P}}{\partial \gamma}\right) \tag{9}$$

Hence, if $\underline{g}$ is the gradient of the negative log-likelihood function, its elements are given by

$$g_i = -\frac{\partial \log\left(f_B\left(\underline{\beta}\right)\right)}{\partial \lambda_i} = \frac{1}{2} trace\left(\mathbf{P}^{-1}\mathbf{P}_i\right) - \frac{1}{2}\underline{\beta}^T \mathbf{P}^{-1}\mathbf{P}_i\mathbf{P}^{-1}\underline{\beta} \tag{10}$$

The Hessian matrix $\mathbf{H}$ has elements given by

$$H_{ij} = -\frac{\partial \log\left(f_B\left(\underline{\beta}\right)\right)}{\partial \lambda_i \partial \lambda_j} = \underline{\beta}^T \mathbf{P}^{-1}\mathbf{P}_i\mathbf{P}^{-1}\mathbf{P}_j\mathbf{P}^{-1}\underline{\beta} - \frac{1}{2} trace\left(\mathbf{P}^{-1}\mathbf{P}_i\mathbf{P}^{-1}\mathbf{P}_j\right) \tag{11}$$

The *Maximum Likelihood* estimate of the elements of $\underline{\lambda}$ can then be found using a Newton iteration

$$\underline{\lambda}_{k+1} = \underline{\lambda}_k - \mathbf{H}_k^{-1}\underline{g}_k \tag{12}$$

with the added bonus that

$$E\left\{\mathbf{H}_k^{-1}\right\} \rightarrow E\left\{\left(\underline{\lambda}_k - E\left\{\underline{\lambda}_k\right\}\right)\left(\underline{\lambda}_k - E\left\{\underline{\lambda}_k\right\}\right)^T\right\} \tag{13}$$

i.e., the expected value of the inverse Hessian is the covariance matrix for the estimates. Of course, the iteration in equation (13) needs an initial estimate. This can be obtained using the other estimation approaches to which we previously alluded. Or, since the short term error is generally dominated by the measurement noise term, the mean square of $\left\{\beta_k \mid k = 1, N\right\}$ is approximately $E\left\{\beta_k^2\right\} \approx 20\sigma_m^2 / T^6$ and, λ_1 can be initialized from this. The other terms can then be initialized to zero. This initialization procedure will probably work most of the time for real data.

One issue that can arise in the iteration defined by equation (12) is that the Hessian matrix as defined in equation (11) is not necessarily positive definite. If, at some step, the Hessian is not positive definite, the iteration will likely diverge. Therefore, it is generally better to use the expected value of the Hessian

$$E\left\{H_{ij}\right\} = \overline{H}_{ij} = \frac{1}{2} trace\left(\mathbf{P}^{-1}\mathbf{P}_i\mathbf{P}^{-1}\mathbf{P}_j\right) \tag{14}$$

rather than equation (11) in the iteration. In addition, for the matrix defined by equation (14) to be positive definite, it is necessary to constrain the iteration so that the stability parameter estimates remain non-negative.

Since $\sqrt{x \pm \varepsilon} \approx \sqrt{x} \pm \dfrac{\varepsilon}{2\sqrt{x}}$ we have the estimate and error bounds for the SGEM parameters as

$$\begin{bmatrix} \hat{\sigma}_m \\ \hat{\sigma}_v \\ \hat{\sigma}_u \end{bmatrix} = \left[\left(\sqrt{\lambda_1} \pm \frac{1}{2}\sqrt{\frac{\overline{H}_{11}}{\lambda_1}} \right) T^3 \quad \left(\sqrt{\lambda_2} \pm \frac{1}{2}\sqrt{\frac{\overline{H}_{22}}{\lambda_2}} \right) T^{2.5} \quad \left(\sqrt{\lambda_3} \pm \frac{1}{2}\sqrt{\frac{\overline{H}_{33}}{\lambda_3}} \right) T^{1.5} \right]^T \qquad (15)$$

This is all fine and good except for computational burdens. Typical data sets from gyro testing are composed of tens of thousands of data points. But, computing e.g., $trace\left(\mathbf{P}^{-1}\mathbf{P}_i\right)$ for such large dimensions is problematic. A single 10,000 × 10,000 matrix using at least 8 bytes for each element requires 800 megabytes of storage. This amount of storage can be significantly cut down by using sparse matrix representations but, the matrix $\mathbf{P}^{-1}\mathbf{P}_i$ is generally not sparse and requires full storage capability. Memory usage can be cut down through additional processing but, only by significantly increasing the processing load, which is already formidable enough for such large matrices.

We are, thus, led to consider suboptimal schemes for estimating the SGEM parameters. One way is to divide the data set up into equal intervals, throwing out the three data points between intervals that would create a cross correlation between them. Then, if there are *K* subsequences of *M* points each, represented by the set of *M*-element vectors $\underline{\beta}_k$, the joint PDF is

$$f_B\left(\left\{\underline{\beta}_k\right\}\right) = \prod_{k=1}^{K} \frac{1}{(2\pi)^{M/2}\sqrt{|P|}} e^{-\frac{1}{2}\underline{\beta}_k^T P^{-1} \underline{\beta}_k} \qquad (16)$$

The elements of the gradient of the negative log-likelihood function are

$$g_i = \frac{K}{2} trace\left(\mathbf{P}^{-1}\mathbf{P}_i\right) - \frac{1}{2}\sum_{k=1}^{K} \underline{\beta}_k^T \mathbf{P}^{-1}\mathbf{P}_i\mathbf{P}^{-1}\underline{\beta}_k \qquad (17)$$

and the expected Hessian matrix elements are

$$\overline{H}_{ij} = \frac{K}{2} trace\left(\mathbf{P}^{-1}\mathbf{P}_i\mathbf{P}^{-1}\mathbf{P}_j\right) \qquad (18)$$

for appropriately sized matrices $\mathbf{P}$ and $\mathbf{P}_{i,j\ldots}$.

The iteration to find a solution proceeds exactly as before in equation (12) using the gradient and Hessian of equations (17) and (18). Clearly, equations (16-18) approach equations (6,10,14) as $M \to N$ ($K \to 1$). We shall refer to this algorithm as the *SGEMML* algorithm.

Consider an example. Assume that average rate data in *rad./sec.* is sampled every 1 second to produce 8192 samples and $\sigma_m = 10^{-3}\, rad$, $\sigma_v = 10^{-4}\, rad / \sqrt{\sec}$, $\sigma_u = 10^{-5}\, rad / \sqrt{\sec^3}$ and $r = 10^{-5}\, rad / \sec^2$. Of course, in the usual circumstance, we do not know these values a priori but, knowing them, we can calculate the expected uncertainty in our estimates versus the length *M* of data segments we choose. We do this by computing the expected Hessian matrix in equation (18) for the actual parameters and using equation (15) to compute the approximate variations in the square roots of the squared parameter estimates. These approximate 1σ error bounds are depicted in Figure 2.

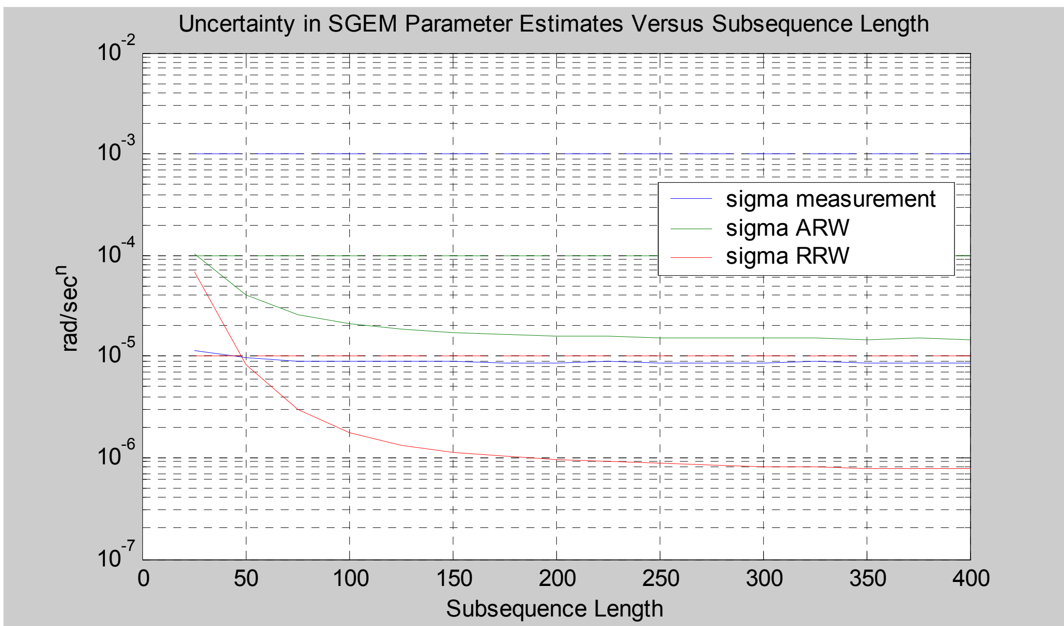

Figure 2 Example SGEM Parameter Uncertainty

As may be seen, the uncertainty begins to level off for the progressively longer term ARW and RRW processes after subsequence length of about *M = 200.*

Usually, estimating a rate ramp is really not of interest because it may not be a repeatable parameter through power cycles and, since it is the value of a random state, you do not need it to implement a Kalman or other type filter. However, attitude filters are often implemented without rate ramp states anyway to reduce processing and, it may be of interest to bound the level of error which can come about due to this neglect. Estimating the ramp can be done optimally without too much of a computational burden using MATLAB's (ref. 7) sparse matrix representations. For this purpose, we use the sequence $\{\alpha_k\}$ in equation (3) expressed as a vector $\underline{\alpha}$. The PDF for this *N+1* point sequence is

$$f_A(\underline{\alpha}) = \frac{1}{(2\pi)^{(N+1)/2}\sqrt{|\mathbf{R}|}} e^{-\frac{1}{2}(\underline{\alpha}-r\underline{u})^T \mathbf{R}^{-1}(\underline{\alpha}-r\underline{u})} \tag{19}$$

where $\underline{u} = [1 \quad 1 \quad 1 \quad \cdots]^T$ is a vector of 1's and $\mathbf{R} = \frac{\sigma_m^2}{T^4}\mathbf{R}_1 + \frac{\sigma_v^2}{T^3}\mathbf{R}_2 + \frac{\sigma_u^2}{T}\mathbf{R}_3$ with

$$\mathbf{R}_1 = \begin{bmatrix} 6 & -4 & 1 & & & \\ -4 & 6 & -4 & 1 & & \\ 1 & -4 & 6 & -4 & 1 & \\ & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & 1 & -4 & 6 & -4 \\ & & & 1 & -4 & 6 \end{bmatrix}$$

$$\mathbf{R}_2 = \begin{bmatrix} 2 & -1 & 0 & & & \\ -1 & 2 & -1 & 0 & & \\ 0 & -1 & 2 & -1 & 0 & \\ & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & 0 & -1 & 2 & -1 \\ & & & 0 & -1 & 2 \end{bmatrix}$$

$$\mathbf{R}_3 = \begin{bmatrix} 2/3 & 1/6 & 0 & & & \\ 1/6 & 2/3 & 1/6 & 0 & & \\ 0 & 1/6 & 2/3 & 1/6 & 0 & \\ & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & 0 & 1/6 & 2/3 & 1/6 \\ & & & 0 & 1/6 & 2/3 \end{bmatrix}$$

Setting the derivative of the negative log-likelihood function to zero, we find the optimal estimate of r to be

$$\hat{r} = \frac{\underline{u}^T \mathbf{R}^{-1} \underline{\alpha}}{\underline{u}^T \mathbf{R}^{-1} \underline{u}} \tag{20}$$

and, the estimate of uncertainty in this value is

$$\sigma_r = \frac{1}{\sqrt{\underline{u}^T \mathbf{R}^{-1} \underline{u}}} \tag{21}$$

This estimate can be computed in MATLAB (ref. 7) using sparse matrix manipulations for **R**.

SIMULATION

This section compares the performance of the AUTOFIT algorithm with the SGEMML algorithm for a particular set of stability parameters. For Monte Carlo comparison, we generated 100 batches of data with 8192 points each and statistics conforming to the SGEM with parameters chosen the same as those which produced Figure 2. The schematic diagram in Figure 3 shows how to generate data with average rate error output autocorrelation function equivalent to that of equation (2) (equal to it with properly defined initial delay states, but this has no effect on the performance of the two estimation algorithms). The first three inputs are independent discrete time white noise processes with the indicated distribution (N for normal, U for uniform) with zero mean and the indicated variance. The fourth input is a deterministic bias plus ramp term.

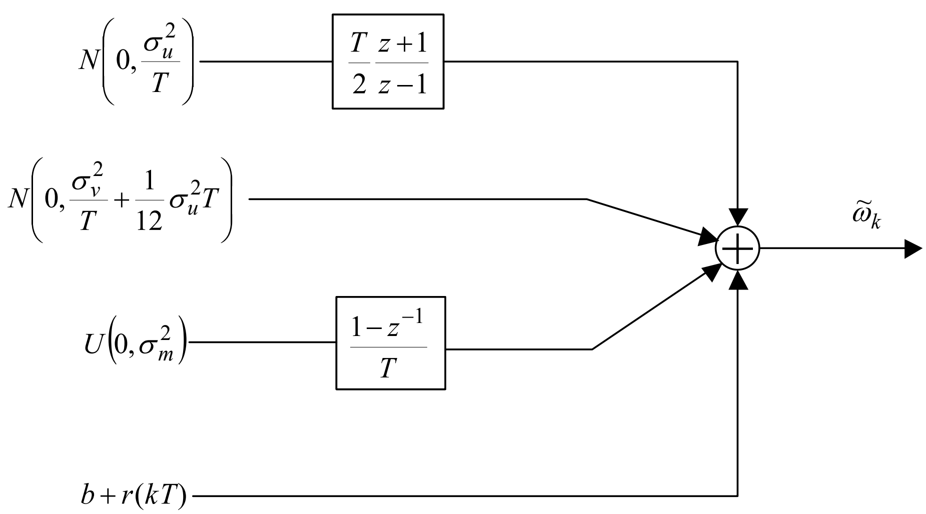

Figure 3. Schematic Diagram of SGEM Data Generator

The histograms in Figure 4 show that, for this particular combination of stability parameters, AUTOFIT and the SGEMML with 200 point subsequences exhibited fairly comparable performance for most of the parameters. A noticeable improvement in performance can be seen in the estimation of the rate random walk parameter where the SGEMML estimates have a narrower spread about the true value.

For all parameters, the predicted 3σ error bounds from the SGEMML look pretty close to bounding the actual results from the SGEMML. These bounds were obtained by multiplying the 1σ bounds in Figure 2 at $M = 200$ by three. In this way, we were able to predict a priori the performance of the SGEMML algorithm in the Monte Carlo runs. In a real world case where we might have one set of gyro data to analyze, the SGEMML automatically computes these error bounds, based on the estimated parameters, as part of the estimation process.

Simulations are excellent for confirming theory but, AUTOFIT has also stood the test of time in demonstrating robust and reliable estimation of stability parameters from real world instruments, some of them significantly affected by environmental disturbances. It remains to be seen if the SGEMML will be as robust to real world data, though there is no reason to expect it would not be. In any case, it would be wise and recommended to supplement such algorithms as these with other tools which give insight into the quality of the underlying model. Sargent and Wyman supplemented AUTOFIT with various charts for examining the long term behavior of the data averages used in the algorithm. We would recommend a PSD analysis of average rate data as being uniquely useful in identifying deviations from the model and sources of disturbances (e.g., disturbances at frequencies associated with heating and cooling cycles in the laboratory, 60 Hz electrical sources, etc...).

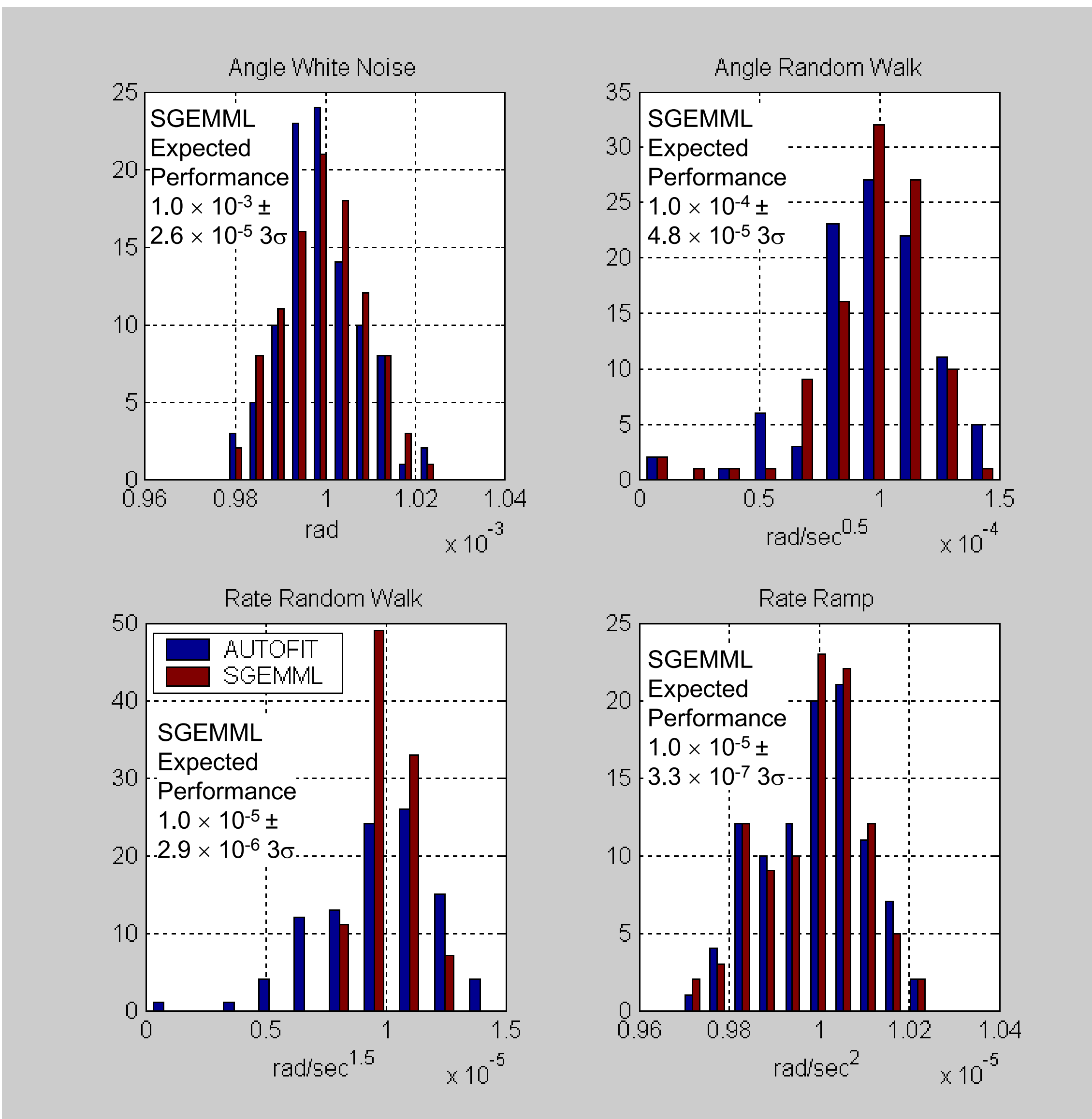

Figure 4. Comparison of AUTOFIT and SGEMML for a Particular Set of Stability Parameters

CONCLUSIONS

The SGEMML algorithm outlined in this paper is probably the most nearly optimal algorithm available for estimating gyroscope stability parameters. A significant advantage of the algorithm is that an estimate of the estimation error is automatically provided.

It may be possible to devise numerical schemes which would make the application of the SGEMML algorithm for longer subsequences of data practical. In any case, the ability to compute the amount of data and subsequence size needed to reach a prescribed level of accuracy is a powerful tool that can be used to plan test conditions to achieve desired results.

Obtaining really good estimates of gyroscope stability parameters is not easy. Gyroscope data collected over a long period of time are often polluted by environmental disturbances and total isolation from these disturbances is often difficult or not feasible to achieve. In such a case, a good way to mitigate environmental disturbances is to test two identical instruments simultaneously so that common disturbances can be subtracted out. Scaling by $1/\sqrt{2}$ then produces a signal which can be considered to be statistically equivalent to a single instrument, at least as far as second order statistics are concerned. Bandlimited disturbances can be filtered out but, it has to be done in such a way that preserves conformance of the data with the SGEM. Other analysis tools such as the PSD can be helpful in identifying corrupted data and determining if and how it can be salvaged.

REFERENCES

1) Markley and Reynolds, "Analytic Steady-State Accuracy of a Spacecraft Attitude Estimator," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 23, no. 6, pp. 1065-1067
2) Oppenheim and Schafer, *Digital Signal Processing.* Upper Saddle River, N.J.: Prentice-Hall, 1974
3) Papoulis, A., *Probability, Random Variables, and Stochastic Processes.* New York: MacGraw-Hill, 2001
4) http://www.allanstime.com/AllanVariance/
5) Sargent and Wyman, "Extraction of Stability Statistics from Integrated Rate Data", 1980 AIAA Guidance & Control Conference, paper #80-1728
6) Mendel, J., *Lessons in Estimation Theory for Signal Processing, Communications, and Control.* Upper Saddle River, N.J.: Pearson Education, 1995
7) http://www.mathworks.com